



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Monitoring for the Datalake

Rizart Dona

CERN

December 9, 2020 - 2nd ESCAPE WP2/DIOS Workshop



Overview

- Monitoring Stack
- Monitoring Architecture
- GFAL Dashboards
- FTS Dashboards
- Rucio Dashboards
- Future Work
- References

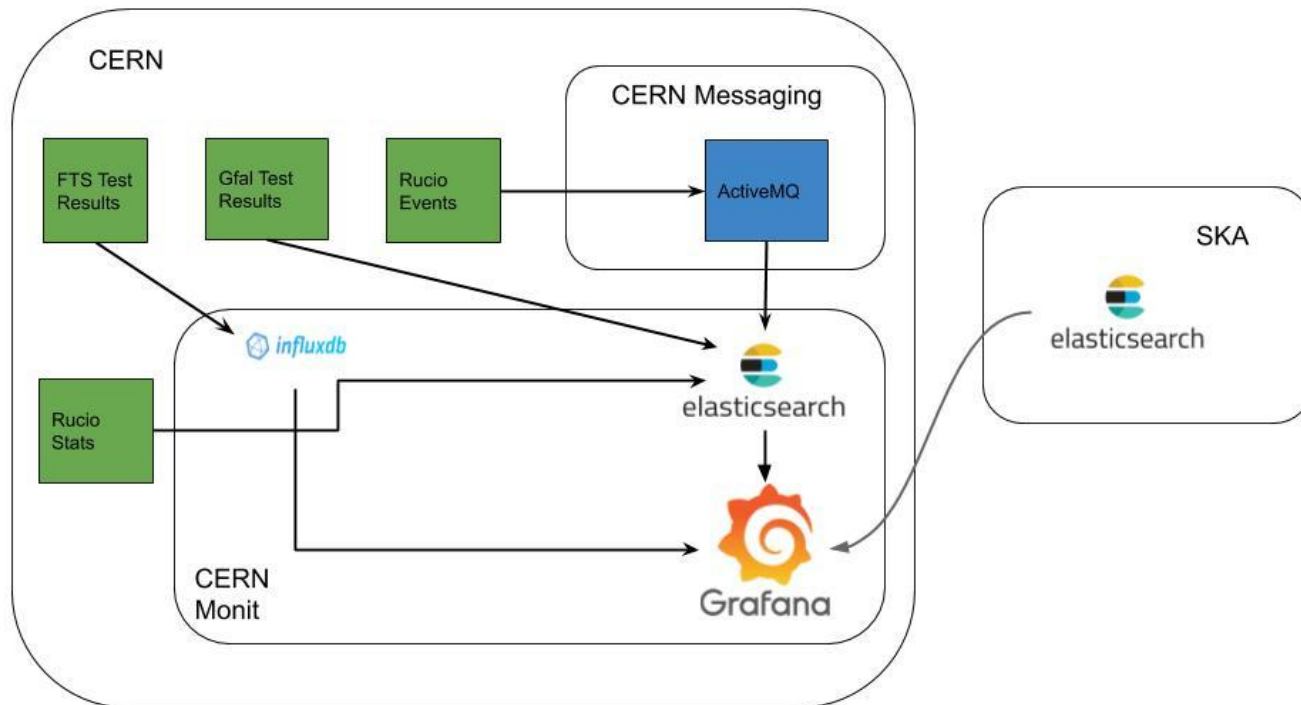


Monitoring Stack

- Backend storage/messaging technologies used
 - **Elasticsearch**, a distributed multitenant-capable full-text search engine
 - **InfluxDB**, an open source time series database that can handle massive volumes and countless sources of time-stamped data
 - **ActiveMQ**, an open source, multi-protocol, Java-based messaging broker
- Visualization platform → **Grafana**
 - A multi-platform open source **analytics** and interactive **visualization** web application
 - Hosted @ CERN by Monit, separate ESCAPE organization
 - Granular permission schema, all interested users already registered
 - Supports multiple datasources (ES, InfluxDB, Graphite, MySQL, etc.)
- Our main datasources
 - **Elasticsearch** @ CERN (hosted by Monit) → Retention policy of 12 months
 - **InfluxDB** @ CERN → Retention policy of 1 month
 - **Elasticsearch** @ SKAO → No retention policy in place, limited by storage given



Monitoring Architecture

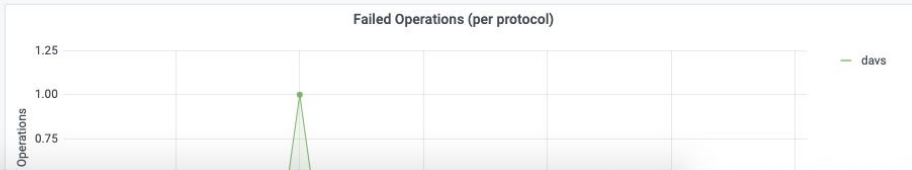
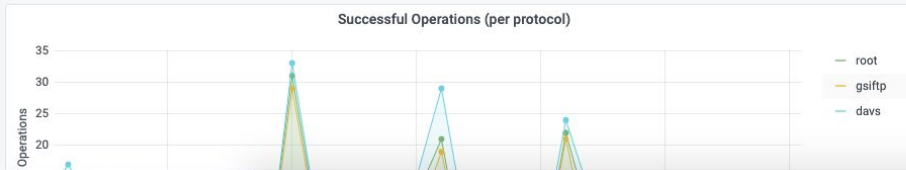
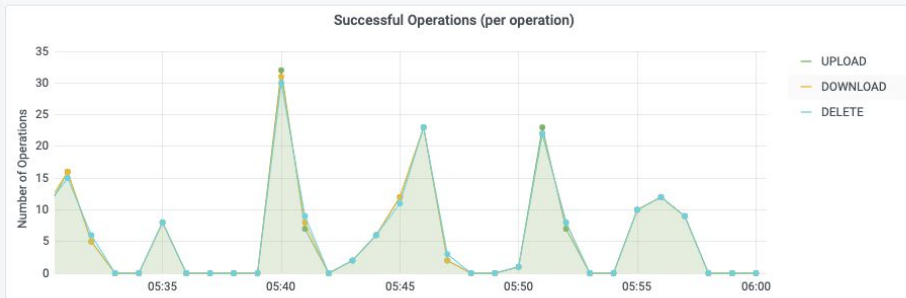
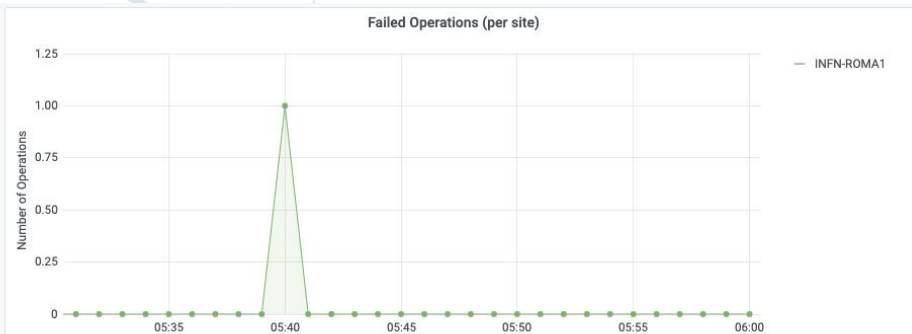
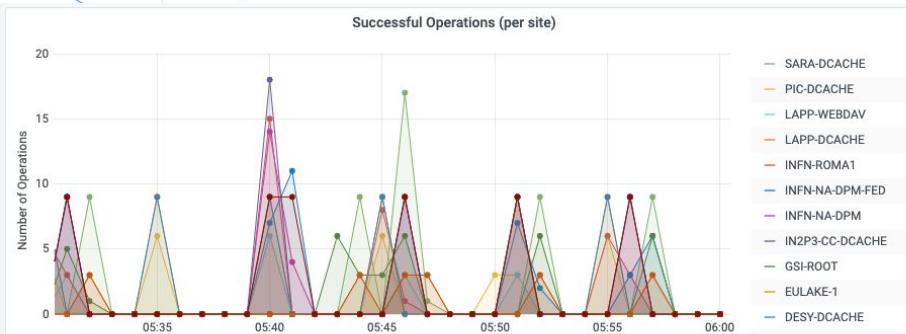


GFAL Dashboards (1/2)

- Data produced and pushed by custom [python script](#)
- Basic GFAL upload/download/delete operations per endpoint
- A user can filter plots by
 - Endpoints
 - Operations (upload, download, delete) - SUCCEEDED/FAILED/SKIPPED
 - Protocol (gsiftp, root, http)
- ES datasource



GFAL Dashboards (2/2)

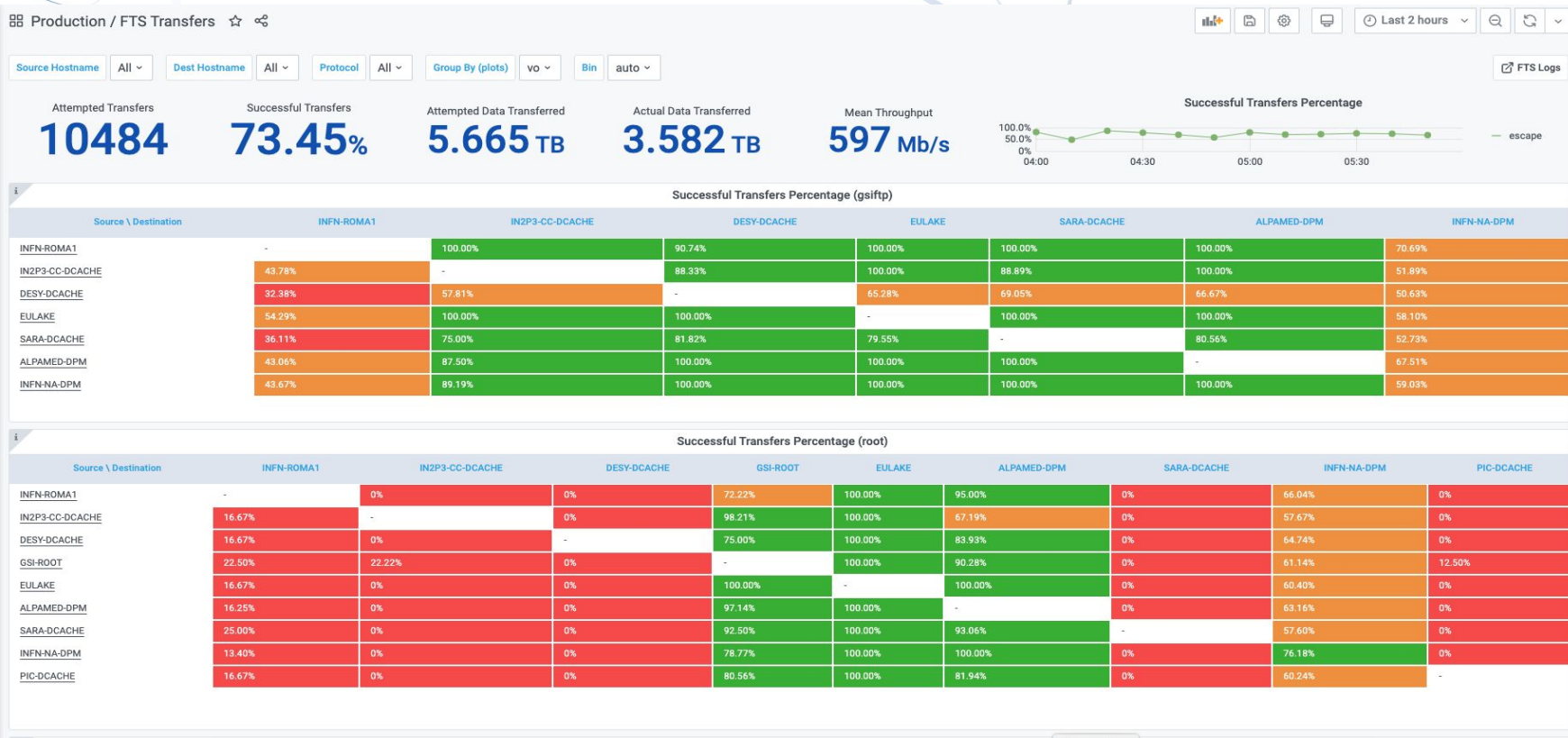


FTS Dashboards (1/3)

- Data produced by the [fts-analysis-datalake](#) toolkit, monitoring basically TPC transfers between the endpoints
- Automatically pushed by the FTS server to ES (configured by FTS team + Monit @ CERN)
- Both InfluxDB + ES datasources
 - Only InfluxDB currently used in the dashboards
 - InfluxDB going to be decommissioned soon → **Migration** of dashboards to ES based needed
- Main highlights for user
 - Aggregated stats (total data transferred, mean throughput, successful transfers, etc.)
 - FTS transfers **efficiency** matrix per protocol
 - Error codes & logfile links
- A lot of tweaking required on the creation aspect to get “beautiful” and intuitive panels
 - e.g. the mapping between endpoints → RSE names was done on the Grafana level



FTS Dashboards (2/3)



FTS Dashboards (3/3)

Production / FTS Transfers ☆ 🔊

Last 2 hours 🔍 ↻

xroot.pic.es	ccdcalltest10.in2p3.fr	https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/a759eb3c-39d8-11eb-9123-fa163e792f12	Error on XrdCl::CopyProcess::Run(): [ERROR] Server responded with an error: [3013] Request 3001 not supported
xroot.pic.es	atlas-dpm-01.roma1.infn.it	https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/1ac51a74-39cf-11eb-8a0a-fa163ece561c	Error on XrdCl::CopyProcess::Run(): [ERROR] Server responded with an error: [3005] destination file prematurely closed
xfer.cr.cnaf.infn.it	t2-dpm-dome.na.infn.it	https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/2d41c95e-39cf-11eb-9c0f-fa163ece561c	TRANSFER ERROR: Copy failed with mode streamed, with error: HTTP 403 : Permission refused
xfer.cr.cnaf.infn.it	dcixwp2dlds1.gsi.de	https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/b5f34d1e-39d8-11eb-8bdf-fa163e792f12	TRANSFER ERROR: Copy failed with mode streamed, with error: (Neon): 404 Not Found
xfer.cr.cnaf.infn.it	atlas-dpm-01.roma1.infn.it	https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/38c600ca-39d8-11eb-8a7a-fa163ec6a1a1	TRANSFER ERROR: Copy failed with mode streamed, with error: HTTP 403 : Permission refused

Transfer Error Codes (per source hostname)

Source Hostname	Error Category	Count ↓
t2-dpm-dome.na.infn.it	OPERATION_NOT_SUPPORTED	236
lapp-testse01.in2p3.fr	OPERATION_NOT_SUPPORTED	127
eoseulake.cern.ch	OPERATION_NOT_SUPPORTED	112
dcache-door-doma01.desy.de	CONNECTION_TIMED_OUT	99
xroot.pic.es	OPERATION_NOT_SUPPORTED	97
atlas-dpm-01.roma1.infn.it	OPERATION_NOT_SUPPORTED	93

Failures By Error Category

Error Category	Count
OPERATION_NOT_SUPPORTED	974
CONNECTION_TIMED_OUT	456
NO_MESSAGE_OF_DESIRED_TYPE	438
PERMISSION_DENIED	160
OPERATION_NOT_PERMITTED	119
HOST_IS_DOWN	98
BAD_ADDRESS	79
INPUT/OUTPUT_ERROR	42
FUNCTION_NOT_IMPLEMENTED	36
INVALID_EXCHANGE	23
COMMUNICATION_ERROR_ON_SEND	16

Transfer Error Codes (per destination hostname)

Destination Hostname	Error Category	Count ↓
ccdcalltest10.in2p3.fr	OPERATION_NOT_SUPPORTED	288
dcache-se-doma.desy.de	OPERATION_NOT_SUPPORTED	245
lobster10.grid.surfsara.nl	OPERATION_NOT_SUPPORTED	229
atlas-dpm-01.roma1.infn.it	NO_MESSAGE_OF_DESIRED_TYPE	227
t2-dpm-dome.na.infn.it	CONNECTION_TIMED_OUT	217
xroot.pic.es	OPERATION_NOT_SUPPORTED	212

Failures By Destination Hostname

Destination Hostname	Count
atlas-dpm-01.roma1.infn.it	592
t2-dpm-dome.na.infn.it	506
ccdcalltest10.in2p3.fr	349
xroot.pic.es	262
lobster10.grid.surfsara.nl	253
dcache-se-doma.desy.de	251
dcixwp2dlds1.gsi.de	97
lapp-testse01.in2p3.fr	41
eoseulake.cern.ch	31
eulakeftp.cern.ch	20
xfer.cr.cnaf.infn.it	19

[Matrices \(3 panels\)](#) |
 [Transfer plots \(2 panels\)](#)

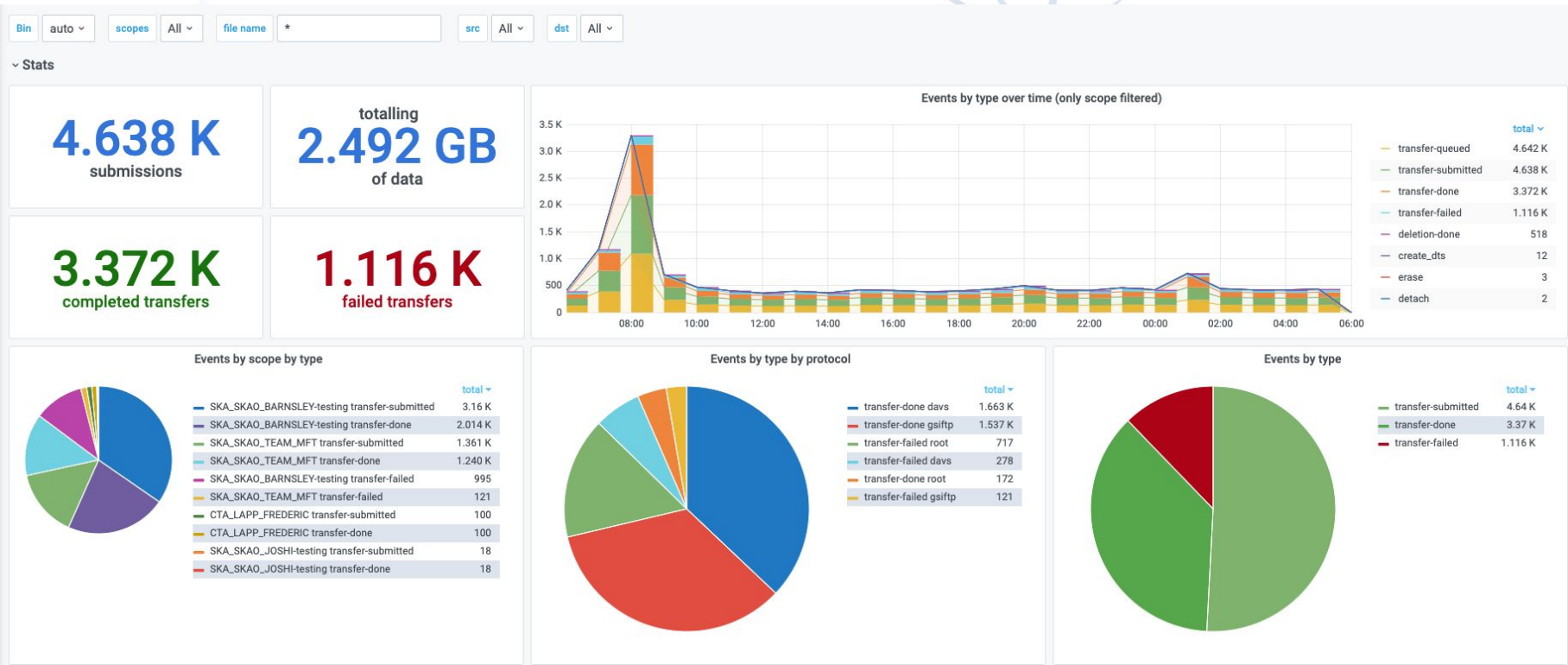


Rucio Dashboards - Rucio Events (1/3)

- Data produced by a Rucio daemon (**Hermes**) → ActiveMQ broker → ES
- Monitoring **replica creation/deletion** (submissions, data transferred, failed transfers, creation of datasets, etc.)
- **Transfer matrix** helps user understand the connectivity between RSES
 - Percentages of successful replication (which are basically reflect into FTS TPC transfers)
 - Recently **optimized**, it feels snappier now (many thanks to Rob Barnsley for helping in this)
- Table with error codes and logfile links to the actual FST transfers



Rucio Dashboards - Rucio Events (2/3)



Rucio Dashboards - Rucio Events (3/3)

src	dst	DESY-DCACHE	SARA-DCACHE	PIC-DCACHE	EULAKE-1	LAPP-DCACHE	IN2P3-CC-	CNAF-STORM	ALPAMED-DPM	GSI-ROOT	INFN-NA-DPM	LAPP-WEBDAV	INFN-NA-DPM-FED	INFN-ROMA1
DESY-DCACHE		NO DATA	67%	60%	100%	100%	0%	100%	100%	100%	100%	100%	NO DATA	NO DATA
SARA-DCACHE		100%	NO DATA	47%	100%	100%	0%	100%	100%	100%	100%	100%	NO DATA	NO DATA
PIC-DCACHE		100%	100%	NO DATA	100%	100%	0%	100%	86%	100%	100%	100%	NO DATA	NO DATA
EULAKE-1		NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA
LAPP-DCACHE		100%	100%	100%	100%	NO DATA	100%	100%	100%	100%	100%	100%	NO DATA	NO DATA
IN2P3-CC-DCACHE		100%	60%	43%	100%	100%	NO DATA	100%	75%	100%	100%	100%	NO DATA	NO DATA
CNAF-STORM		100%	100%	100%	100%	100%	100%	NO DATA	100%	100%	100%	100%	NO DATA	NO DATA
ALPAMED-DPM		100%	86%	91%	100%	100%	0%	100%	NO DATA	100%	100%	100%	NO DATA	NO DATA
GSI-ROOT		100%	100%	100%	100%	100%	100%	100%	100%	NO DATA	100%	100%	NO DATA	NO DATA
INFN-NA-DPM		100%	86%	40%	100%	100%	0%	100%	100%	100%	NO DATA	100%	NO DATA	NO DATA
LAPP-WEBDAV		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	NO DATA	NO DATA	NO DATA
INFN-NA-DPM-FED		NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA
INFN-ROMA1		NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA

Failed Transfers

Failed Transfers								
timestamp ↑	scope	name	size (bytes)	src RSE	dst RSE	protocol	FTS Logfile	failure reason
2020-12-08 23:07:04	SKA_SKAO_BARNESLEY-testing	100KB_081220T22.16.24	100000	ALPAMED-DPM	IN2P3-CC-DCACHE	root	https://fts3-pilot.cer...	TRANSFER [95] Error on XrdCl::CopyProcess::Run(): [ERR
2020-12-08 23:07:50	SKA_SKAO_BARNESLEY-testing	100KB_081220T22.16.39	100000	DESY-DCACHE	PIC-DCACHE	root	https://fts3-pilot.cer...	TRANSFER [5] Error on XrdCl::CopyProcess::Run(): [ERR
2020-12-08 23:12:47	SKA_SKAO_BARNESLEY-testing	100KB_081220T22.16.39	100000	DESY-DCACHE	PIC-DCACHE	root	https://fts3-pilot.cer...	TRANSFER [5] Error on XrdCl::CopyProcess::Run(): [ERR
2020-12-08 23:12:47	SKA_SKAO_BARNESLEY-testing	100KB_081220T22.16.24	100000	ALPAMED-DPM	IN2P3-CC-DCACHE	root	https://fts3-pilot.cer...	TRANSFER [95] Error on XrdCl::CopyProcess::Run(): [ER
2020-12-08 23:12:55	SKA_SKAO_BARNESLEY-testing	100KB_081220T22.16.39	100000	DESY-DCACHE	IN2P3-CC-DCACHE	root	https://fts3-pilot.cer...	TRANSFER [95] Error on XrdCl::CopyProcess::Run(): [ER



Rucio Dashboards - Rucio Statistics (1/3)

- Data source → cron job that fetches periodically (every 10 minutes) the relevant info, pushing to ES
- Novel dashboard, was extensively used during the **FDR** to monitor the progress through the day
- RSE stats
 - Used storage / RSE
 - Number of files / RSE
- Experiment stats
 - Used storage / Experiment
 - DIDs (files, datasets, containers) / Experiment
- Scope stats
 - Used storage / Scope
 - DIDs (files, datasets, containers) / Scope



Rucio Dashboards - Rucio Statistics (2/3)

experiment plots

Used Storage / Experiment (replica=1)



DIDs / Experiment (replica=1)



general stats

NOTE

From this point downwards, this dashboard displays always the latest values of the following statistics. A cron job runs every 10 minutes and registers those values. Timestamp values are UTC.

RSES DATA (all replicas)

122.85 TB

EXPERIMENTS DATA (replica=1)

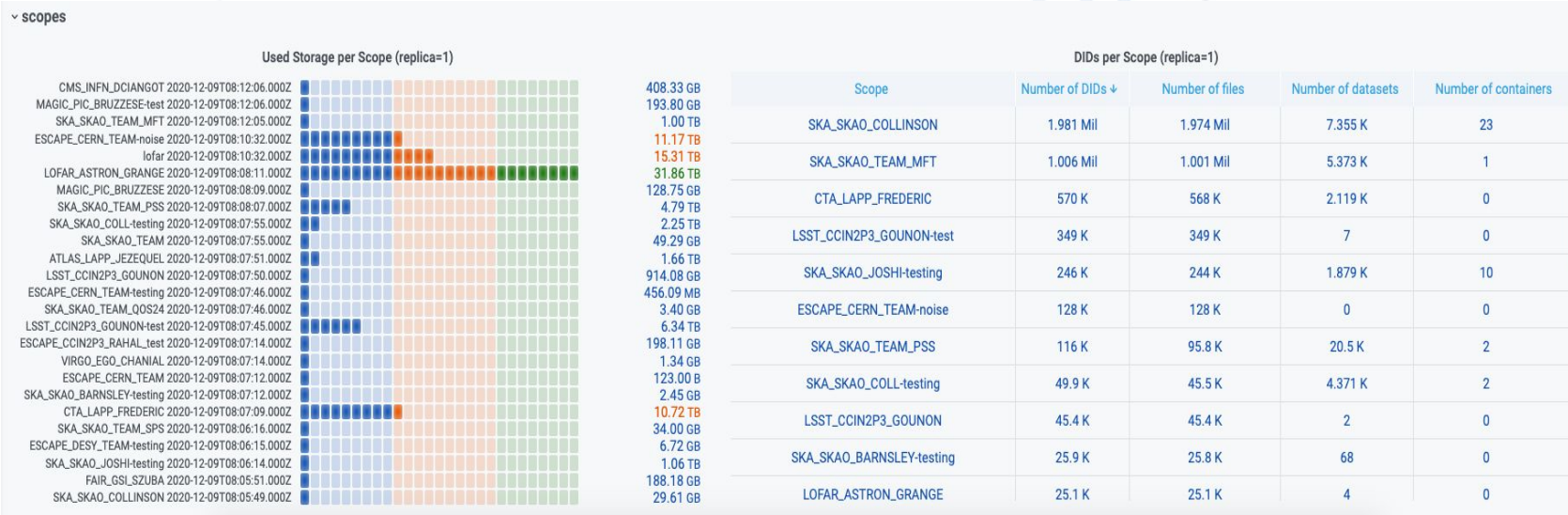
61.63 TB

SCOPES DATA (replica=1)

88.32 TB



Rucio Dashboards - Rucio Statistics (3/3)



Rucio Dashboards - Uploads (1/2)

- Monitoring replica creation + **Rucio uploads** (an insight Rucio Events cannot provide)
- Custom data pushed by SKAO at their ES
- Based on the SKAO Monitoring setup (hosted on STFC Cloud)
 - Elasticsearch deployed on a 3 node bare-metal kubernetes cluster using the helm charts in this [repo](#)

<input type="checkbox"/>	escape-datalake-monitoring-worker-2	Ubuntu-Bionic-NoGui	192.168.200.87	c2.large	rmb-keypair
<input type="checkbox"/>	escape-datalake-monitoring-worker-1	Ubuntu-Bionic-NoGui	192.168.200.31	c2.large	rmb-keypair
<input type="checkbox"/>	escape-datalake-monitoring-head	Ubuntu-Bionic-NoGui	192.168.200.166	c3.large	rmb-keypair

- Persistent Volume provisioner support using [NFS server provisioner](#)
- Largely a dedicated cluster for monitoring at the moment. It is running ES, and any tools needed to enable ES: NFS, Metal load balancer (MetalLB), Ingress controller

```
rjoshi@escape-datalake-monitoring-head:~$ kubectl get pods -n monit
NAME                                READY   STATUS    RESTARTS   AGE
metadata-elasticsearch-coordinating-only-7454647cb5-5z1cm   1/1     Running   0           5d3h
metadata-elasticsearch-coordinating-only-7454647cb5-dtpsm   1/1     Running   0           5d3h
metadata-elasticsearch-data-0                                1/1     Running   0           5d3h
metadata-elasticsearch-data-1                                1/1     Running   0           5d3h
metadata-elasticsearch-master-0                              1/1     Running   0           5d3h
metadata-elasticsearch-master-1                              1/1     Running   0           5d3h
```



Rucio Dashboards - Uploads (2/2)

RSE All ▾

Overview

src	dst	Replication										Upload				
		DESY-DCACHE	SARA-DCACHE	PIC-DCACHE	EULAKE-1	LAPP-DCACHE	IN2P3-CC-DCACHE	CNAF-STORM	ALPAMED-DPM	GSI-ROOT	INFN-NA-DPM	LAPP-WEBDAV	protocol	gsiftp	root	davs
DESY-DCACHE		NO DATA	100%	100%	100%	100%	0%	100%	100%	100%	100%	100%	100%	100%	100%	100%
SARA-DCACHE		100%	NO DATA	100%	100%	100%	0%	100%	100%	100%	100%	100%	100%	100%	100%	100%
PIC-DCACHE		100%	100%	NO DATA	100%	100%	0%	100%	100%	100%	100%	100%	Not Supported	100%	100%	
EULAKE-1		NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	NO DATA	0%	0%	0%	
LAPP-DCACHE		100%	100%	100%	100%	NO DATA	100%	100%	100%	100%	100%	100%	Not Supported	Not Supported	100%	
IN2P3-CC-DCACHE		100%	100%	100%	100%	100%	NO DATA	100%	100%	100%	100%	100%	Not Supported	100%	Not Supported	
CNAF-STORM		100%	100%	100%	100%	100%	100%	NO DATA	100%	100%	100%	100%	Not Supported	Not Supported	100%	
ALPAMED-DPM		100%	100%	83%	100%	100%	0%	100%	NO DATA	100%	100%	100%	100%	100%	100%	
GSI-ROOT		100%	100%	100%	100%	100%	100%	100%	100%	NO DATA	100%	100%	Not Supported	80%	100%	
INFN-NA-DPM		100%	100%	100%	100%	100%	0%	100%	100%	100%	NO DATA	100%	100%	100%	100%	
LAPP-WEBDAV		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	NO DATA	Not Supported	Not Supported	100%	



Future Work

- Migration for FTS dashboard from InfluxDB → ES
- Creation of stable perfSONAR dashboards on Grafana
 - Data is already available
 - Experimental dashboards already exist
 - MadDash currently used as an alternative, need to unify monitoring view
- Interconnect data in Rucio Stats dashboard
- Expose kubernetes cluster monitoring that runs the Rucio daemons



References

- Elasticsearch, <https://www.elastic.co/elasticsearch>
- InfluxDB, <https://www.influxdata.com>
- Grafana, <https://grafana.com>
- Apache ActiveMQ, <http://activemq.apache.org>
- ESCAPE Grafana Org, <https://monit-grafana.cern.ch/d/cHBQ2NiWz/escape-home?orgId=51>



Thank you!
Questions?

